

Expression, purification, cross-reactivity and homology modeling of peanut profilin

Cerrone Cabanos^a, Mary Rose Tandang-Silvas^a, Van Odijk^b, Peter Brostedt^c, Akira Tanaka^d, Shigeru Utsumi^a, Nobuyuki Maruyama^{a,*}

^a Laboratory of Food Quality Design and Development, Graduate School of Agriculture, Kyoto University, Uji, Kyoto 611-0011, Japan

^b Department of Respiratory Medicine and Allergology, Sahlgrenska Academy, Göteborg University, Gothenburg, Sweden

^c Research & Development, Phadia AB, Uppsala, Sweden

^d Phadia KK, Tokyo, Japan

ARTICLE INFO

Article history:

Received 19 February 2010
and in revised form 11 March 2010
Available online 15 March 2010

Keywords:

Peanut allergen
Profilin
Ara h 5
Expression
Cross-reactivity
Homology modeling
Surface-exposed epitopes

ABSTRACT

Plant profilins are known pan-allergens involved in the cross-reactions between pollen and plant foods. Peanut profilin, Ara h 5, is one of the important peanut allergens. Presently, most immunological, biochemical and structural studies on peanut allergens have focused on the three major allergens (Ara h 1, 2 and 3). Here Ara h 5 was cloned, expressed in *Escherichia coli*, Rosetta2(DE3) (Novagen), purified using a combination of ammonium sulfate fractionation and size-exclusion chromatography and yielded a total of 29 mg/l of culture. IgE reactivity was assayed using multiplexed microarray with other peanut allergens (Ara h 1, 2, 3, and 8) and birch (Bet v 2) and timothy (Phl p 2) profilin using sera from peanut allergic Swedish patients. Using homology modeling, Ara h 5 structure was also generated, compared against other profilins and utilized to predict surface-exposed residues potentially forming epitopes. The allergen was recognized by 3 out of 33 sera (9.1%). IgE reactivity to Ara h 5 also coincided with that of two other profilins, Phl p 12 and Bet v 2, confirming cross-reactivity. Interestingly, IgE reactivity to Ara h 5 was higher than above-mentioned profilins which may be indicating specificity of sera towards peanut profilin. Eight surface-exposed epitopes were predicted and verified against experimentally validated sequential epitopes. Three epitopes (#1, 5 and 7) mostly located at the accessible loops and neutral to relatively electropositive sites were found common among profilins, which should be involved in cross-reactivity. A specific putative epitope (#4) was also found which may explain the relative high IgE reactivity to Ara h 5 as compared to the other profilins. Due to its close relation to other allergenic profilins, Ara h 5 could be used as a model and allergen of choice for profilin allergy diagnosis.

© 2010 Elsevier Inc. All rights reserved.

Introduction

Peanut allergy is one of the “big eight” food allergies worldwide accounting for 0.2–1.5% prevalence in children and still rising [1–5]. The rise in peanut allergy is not limited to children. A National Health and Nutritional Examination Survey in the United States (1988–1994) showed that about 8.6% of Americans are sensitized to peanuts [6]. Similarly, national surveys suggest that 1.1% of Americans or three million individuals are allergic to peanuts, tree nuts, or both [3,4]. Symptoms range from mild oral allergy syndrome (OAS)¹ to anaphylactic reactions and even death [7–9]. Thus far, there is no cure for peanut allergy. Therapy focuses mainly on peanut avoidance, early recognition of symptoms due to inadvertent ingestions, and pharmacologic treatment of adverse reactions.

* Corresponding author. Fax: +81 774 38 3761.

E-mail address: marunobu@kais.kyoto-u.ac.jp (N. Maruyama).

¹ Abbreviations used: PDB, protein data bank; OAS, oral allergy syndrome; IPTG, Isopropyl β-D-1-thiogalactopyranoside; IUIS, International Union of Immunological Societies; APBS, adaptive Poisson-Boltzmann solver; RMSD, root mean square deviation; HSE, half sphere exposure; PLP, poly-L-proline.

To date, 11 peanut (*Arachis hypogaea*) allergens have been officially listed by the Allergen Nomenclature Sub-committee of the International Union of Immunological Societies (IUIS) that are termed as Ara h 1 through Ara h 11 (<http://www.allergen.org>). Of these, the three major allergenic peanut proteins are Ara h 1 (63.5 kD), Ara h 2 (17–19 kD) and Ara h 3 (64 kD). All three are storage proteins, Ara h 1 belongs to the vicilin-type 7S globulin while Ara h 2 is a 2S albumin belonging to the conglutin storage proteins. Ara h 3 is an 11S globulin belonging to the glycinin family. Another identified peanut allergen is profilin, which is also known as Ara h 5. Although peanut profilin is considered a minor allergen, it is an important cross-reactive allergen to patients suffering from multivalent type I allergy leading to its designation as a pan-allergen [10]. Profilins are 14–17 kDa proteins which are found in all eukaryotic cells. They were first described by Carlsson and colleagues [11] as actin-binding proteins in calf spleen that are responsible for regulating the normal distribution of filamentous acting networks. Since the discovery of its allergenicity in 1991 [12], profilins were identified as allergen of many fruits, vegetables, foods, as well as pollen. Profilins are highly conserved with

sequence similarity ranging from 70 to 80% among known plant profilins [13].

Most of the immunological, biochemical and structural studies on peanut allergens have focused on the three major allergens. Natural peanut profilin is also difficult to obtain in considerable amount since it is present in very low amounts in crude allergenic extract [28,33]. It has also been shown that several plant profilins can be efficiently produced as stable native-like recombinant proteins in *Escherichia coli* [34] and have been reported to be purified by affinity chromatography either as a histidine-tagged protein (using metal affinity) [35] or as an untagged protein employing poly-L-proline interaction [28,36]. Here we describe the cloning, expression in *E. coli*, and purification of Ara h 5 using an alternative (in the absence of affinity columns) two-step process involving only ammonium sulphate precipitation and size-exclusion chromatography. We also report the IgE reactivity to recombinant Ara h 5 by protein microarray assay of sera from peanut allergic Swedish patients in comparison with other peanut allergens (Ara h 1, 2, 3 and h 8) and known allergenic profilins (Phl p 12 and Bet v 2). We also describe the generated homology structure of Ara h 5, predicted surface-exposed epitopes and compared it against the above-mentioned profilins and two other – Cuc m 2 and Hel a 2 profilin. We also attempted to give insight into the observed relative high reactivity of sera to Ara h 5 against that of other profilins and their cross-reactivity. Complementary to the *in silico* prediction, we further compared the identified potential epitopes of Ara h 5 to the experimentally validated linear or sequential epitopes of other profilins.

Methods

Reverse transcription-polymerase chain reaction (RT-PCR) and cloning of Ara h 5 cDNA

By using RT-PCR, we cloned the cDNA encoding Ara h 5 based from Genbank accession No. **GU354312**. Total RNA was isolated from peanut seeds using TRIzol reagent (Invitrogen). The profilin cDNA was amplified using the RNA LA PCR Kit (AMV), Ver. 1.1 (Takara Bio, Japan). The first-strand cDNA was synthesized with an oligo(dT)-adapter primer containing an M13 primer M4 sequence (contained in the RNA LA PCR Kit). The product was then used for PCR amplification using the forward primer 5'-TCGTGGCAAACCTACGTCGATAACC-3' corresponding to the N-terminus and reverse primer 5'-ATTCAATAAGCTTTAAAGACCCGTATCAATGAGA-3' corresponding to the C-terminal end of the gene containing stop codon (bold font, shown in reverse manner) and *HindIII* restriction site (underlined). This pair of primers was specifically designed to include only the mature region of the Ara h 5 protein and without any polypeptide or protein tag. The reaction was performed using LA Taq DNA Polymerase (Takara Bio) with 30 cycles at 94 °C for 2 min, 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 1 min. The PCR product was further amplified using KOD-Plus DNA Polymerase (TOYOBO, Japan) with 30 cycles at 94 °C for 2 min, 94 °C for 30 s, 55 °C for 30 s, and 68 °C for 1.5 min. The amplified fragment with the expected size was then phosphorylated, and treated with *HindIII*. The resultant fragment was eventually ligated with pET21d (Novagen, CA) that was previously treated with *NcoI*, blunted, treated further with *HindIII* and finally dephosphorylated. Sequencing and insert confirmation was carried out according to the dideoxy method using an ABI Prism 3100 DNA analyzer (Applied Biosystems).

High-level expression and protein purification

The generated pET21d plasmid containing Ara h 5 coding region was transformed into *E. coli*, Rosetta2(DE3) (Novagen, CA) strain.

The recombinant protein was then expressed in LB medium containing 0.5 M NaCl and 50 µg/ml carbenicillin. Initially, three different incubation temperatures were tested, namely, 20, 30 and 37 °C. Although the level of recombinant protein expression was comparable among the three conditions, cells were incubated at 20 °C due the concomitant high expression level of host cell proteins found when incubated at 30 and 37 °C, which may not provide the relative ease in a tagless purification. Cultures were induced by addition of Isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM at OD_{600nm} = 0.60 and incubation continued for 40 h. Cells were harvested by centrifugation at 6700g and pellets were resuspended and sonicated in 35 mM sodium phosphate, pH 7.6, 0.4 M NaCl, 1 mM EDTA, 0.1 mM (*p*-amidinophenyl) methanesulfonyl fluoride (*p*-APMSF), 1.2 µM leupeptin, 0.2 µM pepstatin A, 0.02% (w/v) NaN₃. Insoluble material was removed by centrifugation at 9800g and supernatant was subjected to ammonium sulfate precipitation. Solid ammonium sulfate [(NH₄)₂SO₄] was slowly added to the crude extract up to 35% saturation with stirring on an ice bath. After the sample was kept on ice with stirring for 30 min, it was centrifuged at 4 °C for 30 min (9800g). The supernatant was then further added with ammonium sulfate up to 60% saturation and centrifuged at 9800g. The precipitate was resuspended in the same buffer.

The protein was then applied on HiLoad 16/60 Superdex 75 prep-grade column (GE Healthcare Life Sciences, NJ) equilibrated with the same buffer at a flow rate of 0.5 mL/min. To check for purity, fractions were analyzed by SDS-PAGE using 11% polyacrylamide gels according to the procedure of Laemmli [37], and those containing mainly the bands belonging to Ara h 5 (14.1 kDa) were pooled together as one fraction.

Protein measurement

Protein concentration was determined using a Protein Assay Rapid Kit (Wako, Osaka, Japan) with bovine serum albumin as the standard.

Recombinant Ara h 5 biological activity and protein microarray assay

Peanut allergic Swedish patients (*n* = 33) sera (Phadia AB) selected based on positive case history, skin prick test and ImmunoCAP IgE detection were analysed using multiplexed immunoassay [14]. All allergen reagents were standardized at a concentration of 1 mg/ml and spotted at a volume of 300 µl. Purified Ara h 5 was covalently immobilized to polystyrene particles (0.11 µm diameter, Bangs Laboratories Inc., Carmel, IN, USA) and arrayed onto nitrocellulose membrane together with other peanut allergens (from Phadia AB) – Ara h 1 and 3 (storage proteins), Ara h 2 (2S albumin) and Ara h 8 (pathogenesis-related protein 10, PR-10), and reference profilin allergens Bet v 2 (birch tree) and Phl p 12 (timothy grass). Particles were spotted in an array format using a nano-plotter (Type 2.0 system by Gsellgesellschaft für Silizium-Mikrosysteme mbH, GeSim, Grosserkmannsdorf, Germany). Diameter of arrayed spots was approximately 200 µm and the density of the array was around 400 spots/cm². The assay procedure comprised sequential additions of 30 µl washing buffer (phosphate buffer at neutral pH with blocking agents and detergent), serum (30 µl), washing buffer (30 µl), fluorophore-conjugated monoclonal anti-human IgE antibody (20 µl) and washing buffer (2 × 30 µl), before reading the fluorescence intensity of each spot using a GenePix 4000B laser scanner (Molecular Devices, Sunnyvale, CA, USA). IgE reactivity was reported semi-quantitatively and compared against background binding adjacent to each spot and binding of sera to control spots containing only phosphate-buffered saline solution (PBS).

Homology modeling

In this part of the study, Ara h 5 was compared to either generated model or PDB structure of other profilins, namely; Phl p 12 (Timothy grass, *Phleum pretense*), Bet v 2 (Birch tree, *Betula verrucosa*, 1CQA), Cuc m 2 (Melon, *Cucumis melo*), Hel a 2 (Sunflower, *Helianthus annuus*). The unsolved structures were modeled employing the automated comparative modeling server ModWeb

(www.salilab.org/modweb). In Modweb, models are built for each one of the sequence-structure matches by satisfaction of spatial restraints based on the widely used program Modeller [15]. Ara h 5 (GU354312) and Phl p 12 (CAA54686) were modeled using crystal structure of *Arabidopsis thaliana*, Ara t 8, 3NUL (78% and 76% identity, respectively) as template; Cuc m 2 (AJ565931) using *Hevea brasiliensis*, Heb v 8, 1G5U (83% identity); and Hel a 2 (Y15210) using Bet v 2 (AAA16522), 1CQA (72% identity). Stereochemical

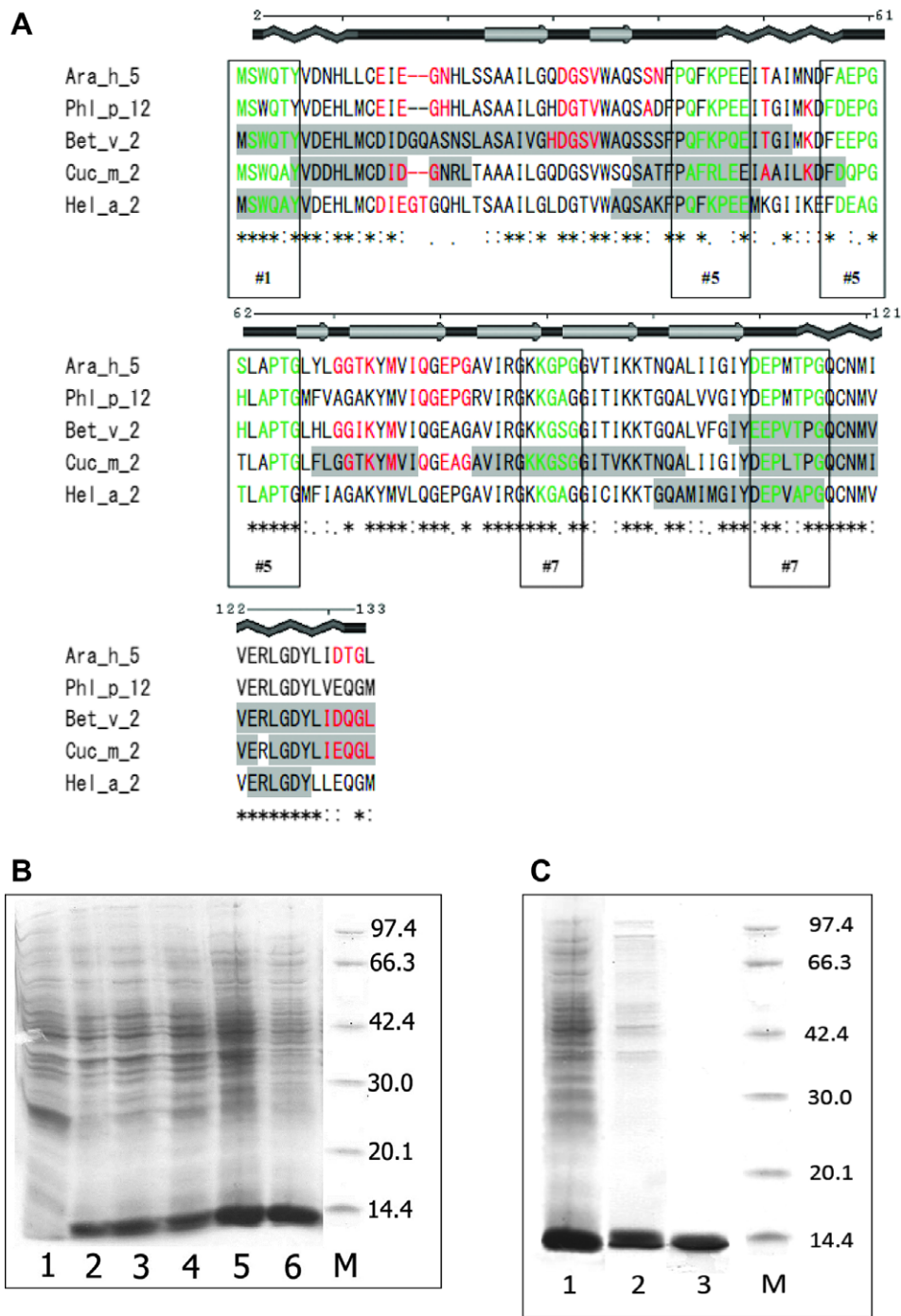


Fig. 1. (A) Amino acid sequence alignment of Ara h 5 (GU354312), Phl p 12 (CAA54686), Bet v 2 (AAA16522), Cuc m 2 (AJ565931) and Hel a 2 (Y15210). Residues in red and green font are the predicted surface-exposed epitopes based on Discotope and BEpro algorithm. Residues in green are the common epitopes of these profilins. On top of the alignment is the consensus secondary structure. — coil or loop, — alpha-helix, and — beta-sheet. (B) Time course of Ara h 5 expression after induction with 1 mM Isopropyl-β-D-thiogalactopyranoside at 20 °C incubation. Lane 1, 0 h; lane 2, 4 h; lane 3, 12 h; lane 4, 24 h; lane 5, 36 h; lane 6, 50 h after induction; lane M, molecular weight markers. (C) Purification of recombinant peanut profilin expressed in Rosetta2(DE3) *E. coli*. Lane 1, soluble lysate; lane 2, after ammonium sulfate fractionation; lane 3, HiLoad Superdex purified recombinant peanut profilin (Ara h 5); lane M, molecular weight markers. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

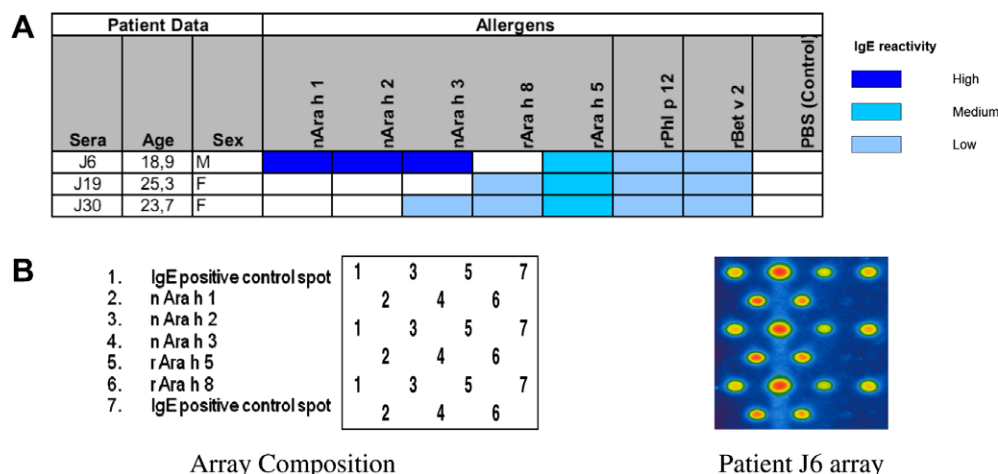


Fig. 2. (A) Individual patient data and summary microarray analysis of Ara h 5 along with Ara h 1, 2, 3, 8, Bet v 2 and Phl p 12 allergens. The low, medium, and high IgE reactivity are shown in different shades of blue. (B) Array plot of patient J6 serum sample. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

quality of the polypeptide backbone and side chains was evaluated using Ramachandran plots generated from RAMPAGE [16]. Amino acid environment of models was evaluated using Verify3D [17].

Prediction of surface-exposed B-cell epitopes

Structure-based prediction of surface-exposed B-cell epitopes was done using the DiscoTope 1.2 [18] and BEpro [19] (formerly PEPITO). Both algorithms combine the amino acid propensity scales and solvent accessibility scores. Discotope uses contact numbers (i.e., the number of C α atoms within a certain distance threshold) while BEpro employs the newly developed half sphere exposure values [20] at multiple distances from target residue in calculating accessibility. Half-sphere exposure (HSE) is a newly developed two-dimensional solvent exposure measure. By conceptually separating an amino acid's sphere in a protein structure into two half spheres which represent its distinct spatial neighborhoods in the upward and downward directions [20]. Both programs generate a pdb-like list file of residues with corresponding score. Here two threshold scores were used for both algorithms. The minimum threshold signifies a sensitivity of 47% and specificity of 75% (−7.7 and 0.9 for Discotope and BEpro, respectively). The upper threshold signifies sensitivity of 32% and a specificity of 85% (−6.0 and 1.1 for Discotope and BEpro, respectively). Residues were then grouped as putative epitopes according to their proximity on surface of the profilins. Structures were superimposed by common geometrical core using Multiprot [21]. Molecular surfaces and electrostatic potentials (APBS plug-in) were calculated and displayed using PyMOL.

Results

Cloning of peanut profilin

The cDNA encoding the allergen Ara h 5 was obtained by using specific primers, containing linker and restriction site, corresponding to the N and C terminus of the mature protein, and was directly cloned into pET21d expression plasmid. The size of PCR product was 406 bp (containing the 396 bp Ara h 5 coding region) at an optimal concentration 1 mM Mg²⁺ and primer annealing at 55 °C. Fig. 1A shows the amino acid sequence alignment of Ara h 5 with the other four profilins: Phl p 12, Bet v 2, Cuc m 2, and Hel a 2. The high sequence similarity is reflected by the presence of 57.9% iden-

tical sites (77 identical residues out of 133) among the five sequences. Ara h 5 sequence is most similar to Cuc m 2 (84.7%), followed by Phl p 12 (78.6%), Bet v 2 (75.2%), and Hel a 2 (69.9%).

Expression and purification of rAra h 5

The final construct was transformed into Rosetta2(DE3) *E. coli*. This strain is a BL21 derivative designed to alleviate codon bias when expressing heterologous proteins in *E. coli*. The cells carry a chloramphenicol-resistant plasmid with the tRNA genes that decode seven rare codons (AGA, AGG, AUA, CUA, GGA, CCC, and CGG) to improve the yield of full-length proteins. The time course of the expression was examined in cell pellet after separation in SDS–PAGE by protein staining with Coomassie blue (Fig. 1B). Major band with an apparent molecular mass of 14.1 kDa can be visualized, and the highest level of production was obtained after 36 h.

A two-step procedure consisting of ammonium sulfate precipitation and size-exclusion chromatography was used to purify recombinant Ara h 5. Cells were harvested by centrifugation, pellet was sonicated, and supernatant was ammonium sulfate fractionated. Fig. 1C shows the pooled fraction of the protein after ammonium sulfate precipitation. The precipitate was then resuspended in the same sonication buffer and loaded on HiLoad 16/60 Superdex 75 pg column. The purified protein (Fig. 1C) migrated as a single band and in agreement with expected size as indicated by gel filtration profile (data not shown) and SDS–PAGE. Total yield of the purification process was 29 mg/l of culture of cells.

Allergenicity and cross-reactivity between Ara h 5 and other profilins

IgE reactivity was further classified into three semi-quantitative categories: low reactivity, medium reactivity, and high reactivity (Fig. 2A). Of the thirty-three peanut allergic Swedish patients, only three sera (J6, J19 and J30) showed significant positive response with medium reactivity to Ara h 5 (9.1%). Serum from patient J6 (Fig. 2B) had high IgE reactivity to Ara 1, 2 and 3. J19 serum had low reactivity to Ara h 8 while J30 had medium reactivity to Ara h 8 and low reactivity to Ara h 3. Ara h 5 also highly coincided with the IgE reactivity of Bet v 2 and Phl p 12 against all sera confirming the well-documented fact of profilin cross-reactivity. Interestingly, IgE reactivity to Ara h 5 appears to be higher than the two profilins which may be indicating higher specificity (i.e. presence of specific epitopes) of sera towards peanut profilin.

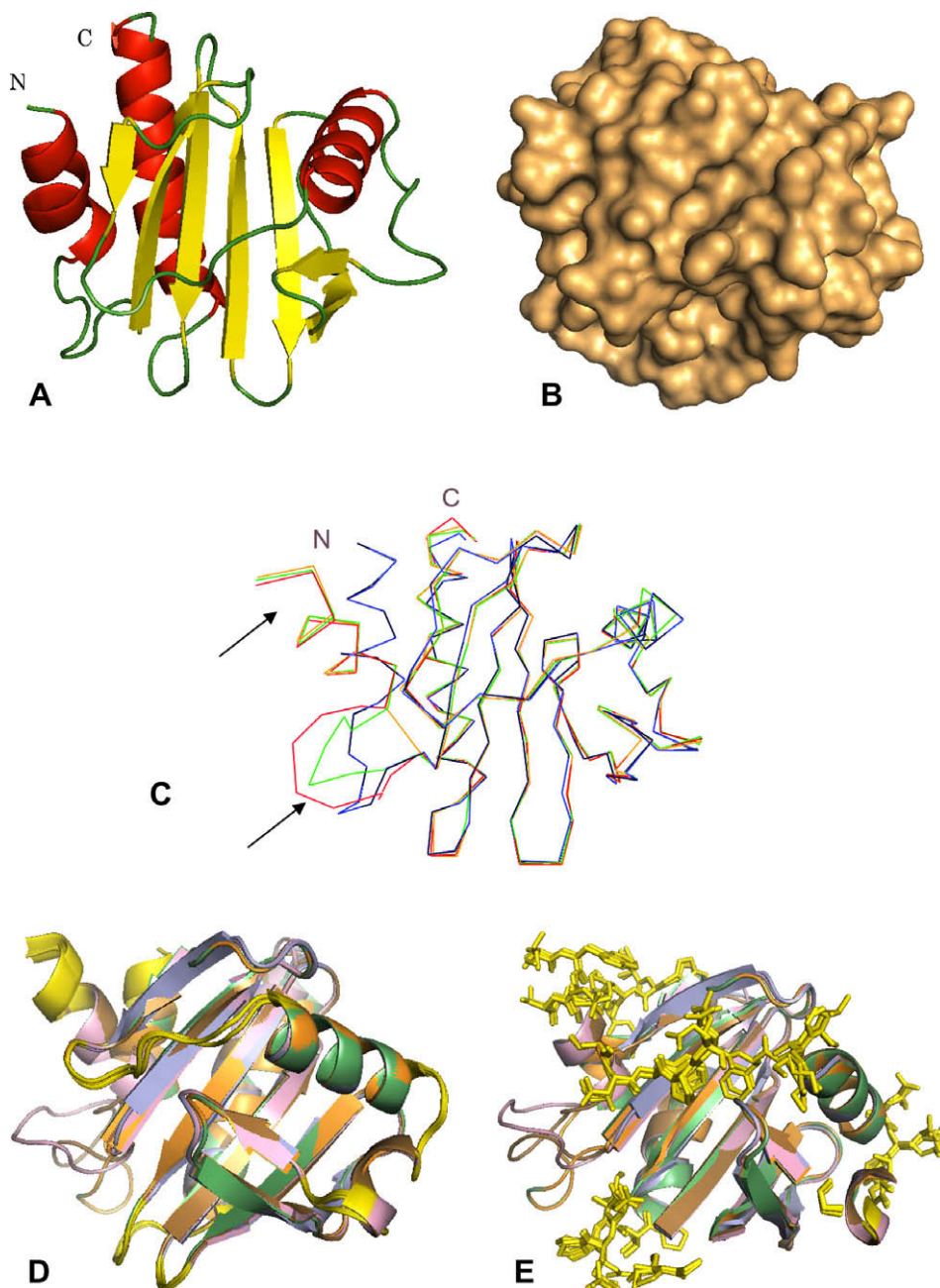


Fig. 3. (A) Ribbon diagram of the three-dimensional model of Ara h 5. Strands of β -sheet and stretches of α -helix are in yellow and red, respectively. Coil structures or loops are in green, N and C indicate the N- and C-terminus of the polypeptide, respectively. (B) Molecular surface of Ara h 5 calculated and displayed with PyMOL. (C) Alpha-carbon superimposition of Ara h 5 (blue), Phl p 12 (black), Bet v 2 (orange), Cuc m 2 (green), and Hel a 2 (red). The top and bottom arrows indicate the N-terminal alpha helix and the variable loop, respectively. (D–E) Ribbon representation of the superimposed structures of Ara h 5, Phl p 12, Bet v 2, Cuc m 2 and Hel a 2 with their consensus epitopes shown in yellow. (D) Coil view and (E) residues involved in the different epitopes are shown as sticks. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Structural analysis and epitope prediction

As antibodies bind to epitopes formed by the spatial arrangement of amino acid residues on surface, model structure of Ara h 5 was generated and compared to model structures of Phl p 12, Cuc m 2 and Hel a 2, and PDB structure of Bet v 2 (1CQA). The structure of Ara h 5 is very identical to the other profilins (Fig. 3A and B). Peanut profilin is composed of three α -helices, seven β -strands and ten turns. It also displays the same motif of seven-stranded but incomplete and anti-parallel up-and-down β -strand barrels observed in other profilins. The similarity between Ara h 5 and Phl p 12 revealed a root mean square deviation (RMSD)

of 0.8 Å while that between Ara h 5 and Bet v 2 showed 1.9 Å. The similarity between Ara h 5 and Cuc m 2 showed an RMSD of 2.1 Å while that between Ara h 5 and Hel a 2 showed a value of 2.48 Å. The only major structural differences as shown in superimposition of their peptide backbones were observed with respect to the direction of the N-terminal α -helix and orientation of the highly accessible variable loop (Fig. 3C–E). Apparently peanut profilin is closest to the backbone structure of timothy grass Phl p 12. Bet v 2 appears to lack the common variable loop conformation for high solvent accessibility. Both Cuc m 2 and Hel a 2 possess a slightly bent loop oriented towards their N-terminal α -helix and may not have the same accessibility as that of Ara h 5 and Phl p 12.

Table 1

Putative surface-exposed epitopes of the five profilins. Their common epitopes are highlighted in gray. Residue number is shown as subscript. Residues in bold font are the amino acids with scores equal to or higher than the upper threshold value (discotope score ≥ -6.0 and z -score ≥ 1.1 for Discotope and BEpro program, respectively) generated by either one or both algorithms. The upper threshold value signifies a detection sensitivity of 32% and a specificity of 85%.

Ara h 5	Phl p 12	Bet v 2	Cuc m 2	Hel a 2
M ₁ S ₂ W ₃ Q ₄ T ₅ Y ₆ (#1)	M ₁ S ₂ Q ₄ T ₅ Y ₆ (#1)	S ₂ W ₃ Q ₄ T ₅ Y ₆ (#1)	M ₁ S ₂ W ₃ Q ₄ Y ₆ (#1)	S ₂ W ₃ Q ₄ A ₅ Y ₆ (#1)
E ₁₄ F ₁₆ G ₁₇ N ₁₈ (#2)	E ₁₄ E ₁₆ G ₁₇ H ₁₈ (#2)		I ₁₅ D ₁₆ G ₁₇ (#2)	D ₁₄ I ₁₅ E ₁₆ G ₁₇ T ₁₈ (#2)
D ₂₉ G ₃₀ S ₃₁ V ₃₂ (#3)	D ₂₉ G ₃₀ T ₃₁ V ₃₂ (#3)	H ₃₀ D ₃₁ G ₃₂ S ₃₃ V ₃₄ (#3)		
S ₃₇ N ₃₈ (#4)				
P ₄₀ Q ₄₁ K ₄₃ P ₄₄ E ₄₅	Q ₄₁ K ₄₃ P ₄₄ E ₄₅ E ₄₆	Q ₄₃ F ₄₄ K ₄₅ P ₄₆ Q ₄₇ E ₄₈	A ₄₁ F ₄₂ R ₄₃ L ₄₄ E ₄₅	Q ₄₃ K ₄₅ P ₄₆ E ₄₇ E ₄₈
T ₄₈ A ₅₅ E ₅₆ P ₅₇ G ₅₈ S ₅₉ P ₆₂ T ₆₃ G ₆₄ (#5)	T ₄₈ K ₅₂ F ₅₄ D ₅₅ E ₅₆ P ₅₇ G ₅₈ H ₅₉ A ₆₁ P ₆₂ T ₆₃ G ₆₄ (#5)	T ₅₀ K ₅₂ E ₅₇ E ₅₈ P ₅₉ G ₆₀ H ₆₁ A ₆₃ P ₆₄ T ₆₅ G ₆₆ (#5)	A ₄₈ K ₅₂ D ₅₅ Q ₅₆ P ₅₇ G ₅₈ H ₅₉ P ₆₂ T ₆₃ G ₆₄ (#5)	D ₅₇ E ₅₈ A ₅₉ G ₆₀ T ₆₁ A ₆₃ P ₆₄ T ₆₅ (#5)
G ₆₈ G ₆₉ T ₇₀ K ₇₁ M ₇₃ L ₇₅ Q ₇₆ E ₇₈ P ₇₉ G ₈₀ (#6)	I ₇₅ Q ₇₆ G ₇₇ E ₇₈ P ₇₉ G ₈₀ (#6)	G ₇₀ G ₇₁ L ₇₂ K ₇₃ M ₇₅ (#6)	G ₆₉ K ₇₁ M ₇₃ Q ₇₆ E ₇₈ A ₇₉ G ₈₀ (#6)	
K ₈₇ G ₈₈ P ₈₉ D ₁₀₇ E ₁₀₈ T ₁₁₁ P ₁₁₂ G ₁₁₃ (#7)	K ₈₇ G ₈₈ A ₈₉ E ₁₀₈ P ₁₀₉ T ₁₁₁ P ₁₁₂ G ₁₁₃ (#7)	K ₈₉ G ₉₀ S ₉₁ G ₉₂ E ₁₀₉ E ₁₁₀ P ₁₁₁ V ₁₁₂ T ₁₁₃ G ₁₁₅ (#7)	K ₈₆ K ₈₇ G ₈₈ S ₈₉ G ₉₀ E ₁₀₈ P ₁₀₉ T ₁₁₁ G ₁₁₃ (#7)	K ₈₈ G ₉₀ A ₉₁ E ₁₁₀ P ₁₁₁ A ₁₁₃ P ₁₁₄ G ₁₁₅ (#7)
D ₁₂₈ T ₁₂₉ G ₁₃₀		I ₁₂₉ D ₁₃₀ Q ₁₃₁ G ₁₃₂ L ₁₃₃ (#8)	I ₁₂₇ E ₁₂₈ Q ₁₂₉ G ₁₃₀ L ₁₃₁ (#8)	

Variations in this loop conformation (loop 1) resulted in the absence of a putative epitope in Bet v 2 (Fig. 1A) and hence have a significant effect on the antigenicity of this region. The difference in antigenicity of this region can be expected as it is highly variable in terms of amino acid sequence among profilins.

Using two algorithms, we found potential surface-exposed sequential and discontinuous epitopes of Ara h 5 and other profilins. Table 1 shows all predicted surface epitopes, which were numbered identically based on Ara h 5, for the five profilins. There were eight epitope sites found in Ara h 5, six sites for Phl p 12, Bet v 2 and Cuc m 2, and four sites for Hel a 2. Four out of eight (#1, 3, 4, and 8) epitopes were sequential in Ara h 5, two out of six (#3 and 6) in Phl p and three out of six (#1, 3 and 8) in Bet v 2. Likewise, three out of six epitopes (#1, 3 and 8) were sequential in Cuc m 2 and two out of four in Hel a 2 (#1 and 2). Most portions of these epitopes, depicted as yellow and green (consensus epitopes) colored surface patches, were located on the front surface of each profilin (Fig. 4A–E). We also mapped the electrostatic potentials on the surface of profilin based on the pdb structure of Bet v 2 and found that most patches correspond to the highly-exposed and relatively neutral to electropositive (Lys-rich) regions which were on the front side (Fig. 4F).

There were also some epitopes found common to some but not in all five profilins. Epitopes #2 was common to all except Bet v 2, while epitope #3 was common only to Ara h 5, Phl p 12 and Bet v 2. Epitopes #6 was found in all except Hel a 2 and epitope #8 was found only in Ara h 5, Bet v 2 and Cuc m 2. Epitope #4 (S₃₇N₃₈) was present only in Ara h 5, which remarkably has the most epitopes encompassing the entire protein. This could explain the relatively high IgE reactivity in Fig. 2. As expected from structure similarity, Ara h 5 shared more epitopes with Phl p 12 (#1–3, 5–7), Bet v 2 (#1, 3, 5–8) and Cuc m 2 (#1, 2, 5–8) than with Hel a 2 (#1, 2, 5, 7). Three epitopes, also on the front side, were found highly similar in sequence and conformation in all five proteins. These epitopes should be the regions responsible for cross-reactivity. Based on Ara h 5 sequence, these were the M₁S₂W₃Q₄T₅Y₆ (#1), P₄₀Q₄₁K₄₃P₄₄E₄₅T₄₈A₅₅E₅₆P₅₇G₅₈S₅₉P₆₂T₆₃G₆₄ (#5), and K₈₇G₈₈P₈₉G₉₀D₁₀₇E₁₀₈P₁₀₉T₁₁₁P₁₁₂G₁₁₃ (#7). Residues in bold typeface are the ones which had scores equal to or higher than the upper threshold score (−6.0 and 1.1 for Discotope and BEpro, respectively), which may also mean high antigenicity. Epitope #1 is composed mostly of neutral to weakly polar residues (M1, S2, W3, T5, Y6) and one neutral residue (Q3). Epitope #5 is composed of 10 neutral to weakly polar residues (P₄₀Q₄₁T₄₈A₅₅P₅₇G₅₈S₅₉P₆₂T₆₃G₆₄) and four polar residues (Q₄₁K₄₃E₄₅E₅₆). Epitope #7 is composed of six neutral to weakly polar (G₈₈P₈₉G₉₀P₁₀₉T₁₁₁) and three polar residues (K₈₇D₁₀₇E₁₀₈). Fig. 5 revealed in a close-up view the conformational similarity of epitopes #1, 5, and 7 among profilins. Most pronounced resemblance was shown between Ara h 5 and Phl p 12. We further looked at the common secondary structural features of these consensus epitopes. Superimposition of ribbon structures (Fig. 3D–E, Fig. 1A) showed that residues were located mostly at the highly accessible loop regions of profilins. Epitope #1 is an exception which was located at the N-terminal site of α -helix 1. Most residues of epitope #5 were found at the loops flanking α -helix 2 and a few residues within the helix itself. Portions of epitope #7 were located at the loops in between β -sheets 5 and 6, and in between β -sheet 7 and α -helix 3. Fig. 3E depicts the stick view of their residues forming a protruding interface.

Discussion

Profilin is an ubiquitous protein found in almost all eukaryotic cells. It was first identified as a minor allergen in birch pollen. So far, profilins were described as pan-allergens in botanically closely

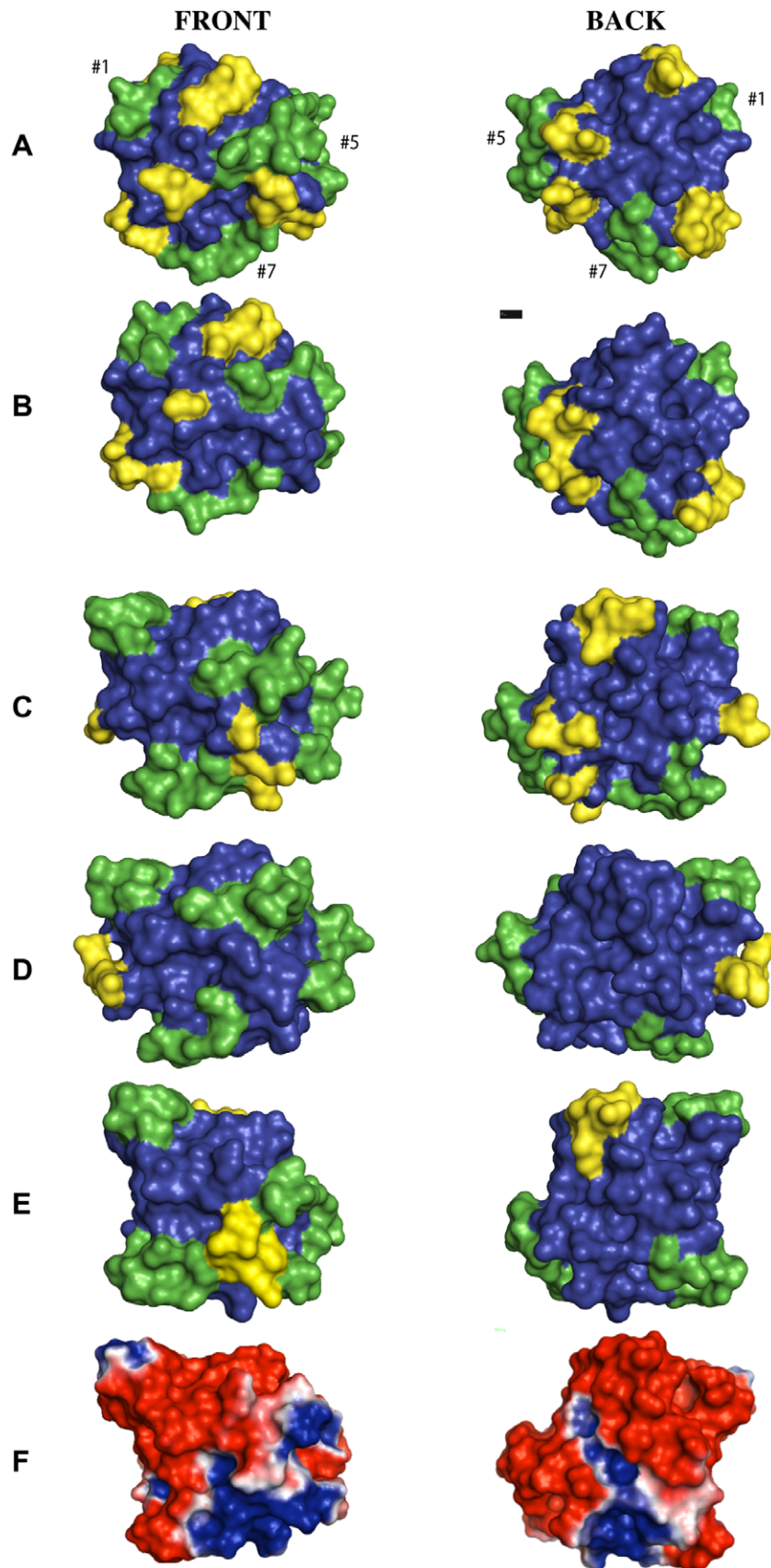


Fig. 4. Putative exposed epitopes (in both yellow and green colored patches) of Ara h 5 (A), Phl p 12 (B), Cuc m 2 (C), Hel a 2 (D), and Bet v 2 (E) mapped on their surfaces (front and back side). Epitope regions in green (#1, 5 and 7) are the shared or consensus epitopes of the three profilins. (F) Mapping of electrostatic potentials, calculated with APBS of PyMOL, on the molecular surface of Bet v 2 (front and back side). Electropositive and electronegative regions are in blue and red, respectively. Neutral regions are colored white. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

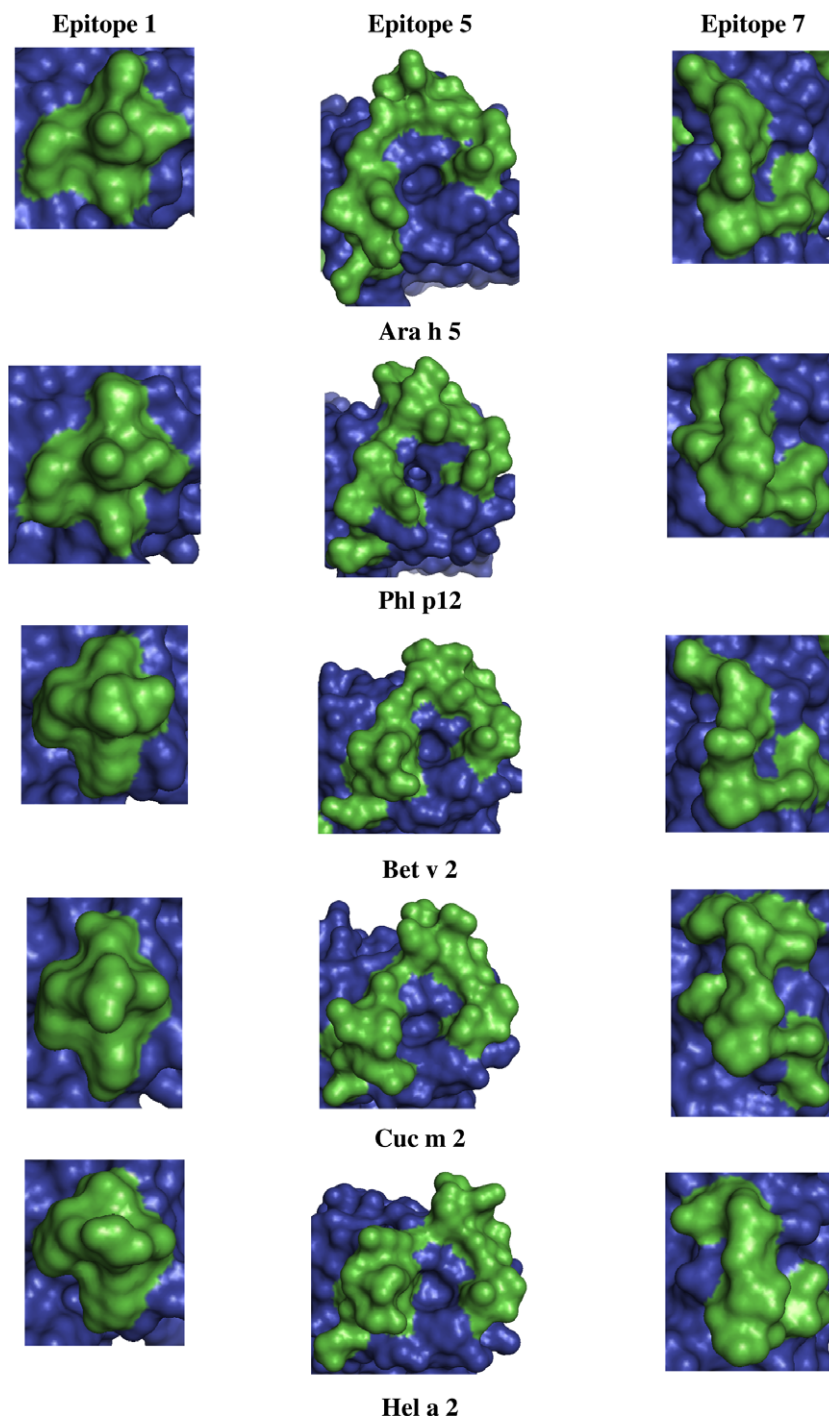


Fig. 5. Three-dimensional conformation of the B-cell epitopes #1, 5 and 7 (green patches) rendered on the molecular surfaces of the five profilins. These images were calculated and displayed with PyMOL. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

related and distantly related plants or unrelated plant [12,13]. This study is primarily aimed at investigating the role of the pan-allergen profilin as a cross-reactive allergen in peanut and giving insight into its cross-reactivity by structural homology modeling. In doing so, we cloned the cDNA encoding Ara h 5 and expressed it in Rosetta2(DE3) (Novagen) *E. coli* strain. Previous initial studies have mentioned the use of BL21(DE3) and recently BL21-Codon-Plus(DE3)-RIL, containing extra copies of tRNA genes decoding for rare codons in *E. coli*, to compensate for any incompatibility between the codon preference in the mRNA of the recombinant gene and in the *E. coli* host [38]. The strain used in this study is also a

BL21 derivative designed to alleviate codon bias when expressing heterologous proteins in *E. coli*. This superior strain carries a chloramphenicol-resistant plasmid with the tRNA genes that decode seven rare codons (AGA, AGG, AUA, CUA, GGA, CCC, and CGG) to improve the yield of full-length proteins. We were also able to purify the protein using a two-step process, consisting of ammonium sulfate fractionation and size-exclusion chromatography, and yielded a total of 29 mg/l of cell culture, which is comparable with the yield of affinity chromatography-based (either using Poly-L-proline-Sepharose or Ni-NTA resin) peanut profilin purifications [28,35,38]. Our procedure can also be a good alternative option

in the absence of a suitable affinity column or resin. In contrast with PLP-Sepharose-based affinity chromatography where high concentration of urea is used to elute bound profilin [28], our procedure uses non-denaturing reagents and eliminates the need for long period of dialysis (at least 24 h) to remove urea from the purified fractions.

Using the purified allergen, we were able to detect IgE reactivity of peanut allergic Swedish patients' sera to peanut profilin employing multiplexed protein microarray assay. This highly sensitive and high-throughput, though semi-quantitative, immunoassay makes it possible to determine specific IgE against multiple allergens simultaneously in one same patient, with a minimum amount of serum [22]. Here we presented that this technique can also aid in detecting cross-reactions. Three (J6, J19 and J30) of thirty-three sera (9.1%) had IgE reactivity to the recombinant profilin. We compared the reactivity profile of Ara h 5 with that of other peanut allergens (Ara h 1, 2, 3 and 8) (data not shown) and noted that there seems to be two groups of individuals in terms of reactivity to peanut allergens. As can also be inferred from previous studies [23–29], the first group is the high response group that appears to react strongly with the three major allergens (Ara h 1, 2 and 3) and probably their homologous allergens (Ara h 4 and 6). The fact that these major allergens are present in large amount in peanuts [24–26] could mean that this group can be a high risk group. The second group is the low to medium response group comprised of individuals who may not necessarily react to major allergens but reactive to the minor allergens, such as, Ara h 5 and 8. We may consider this group as a low risk group. It seems that patients J6, J19, and J30 belong to the latter group. Furthermore, Ara h 5 reactivity pattern showed strong correspondence to other profilins (Phl p 12 and Bet v 2) indicating cross-reactivity. The two other profilins gave positive reaction also only in the three patient sera. The presence of very high similarity in terms of RMSD (ranging from 0.8 to 2.4 Å) among model structures and known pdb structure of profilin signifies that the generated structures using homology modeling are reliable enough for structure-based epitope prediction. Eight surface-exposed epitopes were found in Ara h 5 namely; M₁S₂W₃-Q₄T₅Y₆ (#1), E₁₄E₁₆G₁₇N₁₈ (#2), D₂₉G₃₀S₃₁V₃₂ (#3), S₃₇N₃₈ (#4), P₄₀Q₄₁K₄₃P₄₄E₄₅T₄₈A₅₅E₅₆P₅₇ G₅₈S₅₉P₆₂T₆₃G₆₄ (#5), G₆₈G₆₉T₇₀-K₇₁M₇₃L₇₅Q₇₆E₇₈P₇₉G₈₀ (#6), K₈₇G₈₈P₈₉G₉₀D₁₀₇E₁₀₈P₁₀₉T₁₁₁P₁₁₂-G₁₁₃ (#7) and D₁₂₈T₁₂₉G₁₃₀ (#8). These putative epitopes, except for some portions of epitope #5 and 6, strongly coincided with the regions in Bet v 2, Cuc m 2 and Hel a 2 containing experimentally validated linear epitopes (shaded regions in Fig. 1A) [30–32]. This strong correspondence between the predicted epitopes and experimentally-determined linear epitopes further affirms our confidence in the two algorithms (Discotope and BEpro) used.

Moreover, three epitopes (#1, 5, and 7) were shown to be very similar in all the five profilins in terms of sequence and three-dimensional conformation (Table 1, Fig. 5). Residues of these three epitopes were also mostly located at the loop regions and neutral to relatively electropositive sites on the protein surface. This seems to be expected since the loop regions are also the most accessible site in terms of solvent exposure in profilins. We also noticed that epitopes #5 and 7 also partially corresponded to Lys-rich actin-binding site of profilin, which explains the reported partial inhibition of IgE reactivity to the protein upon treatment with actin [30]. Interestingly, the linear epitopes (shaded regions in Fig. 1A) overlapping epitope #1, portions containing epitope # 5 and 7 are the three highly IgE reactive epitopes in sunflower profilin, Hel a 2 [31]. Among these three, the linear epitope containing portion of #5 is the most reactive epitope in Hel a 2 as well as in Cuc m 2 [32]. All or any of these three epitopes should therefore be involved in the cross-reactivity between different profilins. Remarkably, Ara h 5 also possessed most epitopes found in the entire length of profilins and one specific epitope #4 (S₃₇N₃₈), which

might explain the higher IgE reactivity of sera to Ara h 5 than that of Phl p 12 and Bet v 2.

This study shows that the interplay of common and unique epitopes in influencing the IgE reactivity to Ara h 5 and, in general, profilin allergenicity and cross-reactivity. All these point to the plausibility of using Ara h 5 as a model and allergen of choice for profilin allergy diagnosis.

References

- [1] J. Grundy, S. Matthews, B. Bateman, T. Dean, S.H. Arshad, Rising prevalence of allergy to peanut in children: data from 2 sequential cohorts, *J. Allergy Clin. Immunol.* 110 (2002) 784–789.
- [2] M. Osterballe, T.K. Hansen, C.G. Mortz, A. Host, C. Bindslev-Jensen, The prevalence of food hypersensitivity in an unselected population of children and adults, *Pediatr. Allergy Immunol.* 16 (2005) 567–573.
- [3] S.H. Sicherer, A. Munoz-Furlong, A.W. Burks, H.A. Sampson, Prevalence of peanut and tree nut allergy in the US determined by a random digit dial telephone survey, *J. Allergy Clin. Immunol.* 103 (1999) 559–562.
- [4] S.H. Sicherer, A. Munoz-Furlong, H.A. Sampson, Prevalence of peanut and tree nut allergy in the United States determined by means of a random digit dial telephone survey: a 5-year follow-up study, *J. Allergy Clin. Immunol.* 112 (2003) 1203–1207.
- [5] S.E. Emmett, F.J. Angus, J.S. Fry, P.N. Lee, Perceived prevalence of peanut allergy in Great Britain and its association with other atopic conditions and with peanut allergy in other household members, *Allergy* 54 (1999) 380–385.
- [6] S.J. Arbes Jr, P.J. Gergen, L. Elliott, D.C. Zeldin, Prevalences of positive skin test responses to 10 common allergens in the US population: results from the third National Health and Nutrition Examination Survey, *J. Allergy Clin. Immunol.* 116 (2005) 377–383.
- [7] S.A. Bock, A. Munoz-Furlong, H.A. Sampson, Fatalities due to anaphylactic reactions to foods, *J. Allergy Clin. Immunol.* 107 (2001) 191–193.
- [8] G. Roberts, Anaphylaxis to foods, *Pediatr. Allergy Immunol.* 18 (2007) 543–548.
- [9] H.S. Skolnick, M.K. Conover-Walker, C.B. Koerner, H.A. Sampson, A.W. Burks, R.A. Wood, The natural history of peanut allergy, *J. Allergy Clin. Immunol.* 107 (2001) 367–374.
- [10] R. Valenta, M. Duchene, C. Ebner, P. Valent, C. Sillaber, P. Deviller, F. Ferreira, M. Tejkl, H. Edelmann, D. Kraft, Profilins constitute a novel family of functional plant allergens, *J. Exp. Med.* 175 (1992) 377–385.
- [11] L. Carlsson, L.E. Nystrom, I. Sundkvist, F. Markey, U. Lindberg, Actin polymerizability is influenced by profilin, a low molecular weight protein in non-muscle cells, *J. Mol. Biol.* 115 (1977) 465–483.
- [12] R. Valenta, M. Duchene, K. Pettenburger, C. Sillaber, P. Valent, P. Bettelheim, M. Breitenbach, H. Rumpold, D. Kraft, O. Scheiner, Identification of profilin as a novel pollen allergen; IgE autoreactivity in sensitized individuals, *Science* 253 (1991) 557–560.
- [13] C. Radauer, K. Hoffmann-Sommergruber, Profilins, in: C. Mills, P. Shewry (Eds.), *Plant Food Allergens*, Blackwell, Oxford, UK, 2004, pp. 105–124.
- [14] A. Adachi, T. Horikawa, H. Shimizu, Y. Sarayama, T. Ogawa, S. Sjölander, A. Tanaka, T. Moriyama, Soybean β -conglycinin as the main allergen in a patient with food-dependent exercise-induced anaphylaxis by tofu: food processing alters pepsin resistance, *Clin. Exp. Allergy* 39 (2009) 167–173.
- [15] A. Sali, T.L. Blundell, Comparative modelling by satisfaction of spatial restraints, *J. Mol. Biol.* 234 (1993) 779–815.
- [16] S.C. Lovell, I.W. Davis, W.B. Arendall III, P.I.W. de Bakker, J.M. Word, M.G. Prisant, J.S. Richardson, D.C. Richardson, Structure validation by C alpha geometry: phi, psi and C beta deviation, *Protein Struct. Funct. Genet.* 50 (2002) 437–450.
- [17] R. Luthy, J.U. Bowie, D. Eisenberg, Assessment of protein models with three-dimensional profiles, *Nature* 356 (1992) 83–85.
- [18] P.H. Andersen, M. Nielsen, O. Lund, Prediction of residues in discontinuous B cell epitopes using protein 3D structures, *Protein Sci.* 15 (2006) 2558–2567.
- [19] M.J. Sweredoski, P. Baldi, PEPITO: improved discontinuous B-cell epitope prediction using multiple distance thresholds and half sphere exposure, *Bioinformatics* 24 (2008) 1459–1460.
- [20] T. Hamelryck, An amino acid has two sides: a new 2D measure provides a different view of solvent exposure, *Proteins Struct. Funct. Bioinf.* 59 (2005) 38–48.
- [21] M. Shatsky, R. Nussinov, H.J. Wolfson, A method for simultaneous alignment of multiple protein structures, *Proteins Struct. Funct. Bioinf.* 56 (2004) 143–156.
- [22] M. Ferrer, M.L. Sanz, J. Sastre, J. Bartra, A. del Cuvillo, J. Montoro, I. Jáuregui, I. Dávila, J. Mullol, A. Valero, Molecular diagnosis in allergology: application of the microarray technique, *J. Invest. Allergol. Clin. Immunol.* 19 (Suppl. 1) (2009) 19–24.
- [23] F. Blanc, K. Adel-Patient, M.F. Drumare, E. Paty, J.M. Wal, H. Bernard, Capacity of purified peanut allergens to induce degranulation in a functional in vitro assay: Ara h 2 and Ara h 6 are the most efficient elicitors, *Clin. Exp. Allergy* 39 (2009) 1277–1285.
- [24] A.W. Burks, Peanut allergy, *Lancet* 371 (2008) 1538–1546.
- [25] A.E. Flinterman, E. van Hoffen, C.F. den Hartog Jager, S. Koppelman, S.G. Pasmans, M.O. Hoekstra, C.A. Bruijnzeel-Koomen, A.C. Knulst, E.F. Knol,

- Children with peanut allergy recognize predominantly Ara h2 and Ara h6, which remains stable over time, *Clin. Exp. Allergy* 37 (2007) 1221–1228.
- [26] S.J. Koppelman, M. Wensing, M. Ertmann, A.C. Knulst, E.F. Knol, Relevance of Ara h1, Ara h2 and Ara h3 in peanut-allergic patients, as determined by immunoglobulin E Western blotting, basophil-histamine release and intracutaneous testing: Ara h2 is the most important peanut allergen, *Clin. Exp. Allergy* 34 (2004) 583–590.
- [27] G.W. Palmer, D.A. Dibbern Jr., A.W. Burks, G.A. Bannon, S.A. Bock, H.S. Porterfield, R.A. McDermott, S.C. Dreskin, Comparative potency of Ara h 1 and Ara h 2 in immunochemical and functional assays of allergenicity, *Clin. Immunol.* 115 (2005) 302–312.
- [28] T. Kleber-Janke, R. Crameri, S. Scheurer, S. Vieths, W.M. Becker, Patient-tailored cloning of allergens by phage display: peanut (*Arachis hypogaea*) profilin, a food allergen derived from a rare mRNA, *J. Chromatogr. B Biomed. Sci. Appl.* 756 (2001) 295–305.
- [29] A. Mari, E. Scala, P. Palazzo, S. Ridolfi, D. Zennaro, G. Carabella, Bioinformatics applied to allergy: allergen databases, from collecting sequence information to data integration. The allergome platform as a model, *Cell. Immunol.* 244 (2006) 97–100.
- [30] A.A. Fedorov, T. Ball, N.M. Mahoney, R. Valenta, S.C. Almo, The molecular basis for allergen cross-reactivity: crystal structure and IgE-epitope mapping of birch pollen profilin, *Structure* 5 (1997) 33–45.
- [31] J.A. Asturias, N. Gómez-Bayón, M.C. Arilla, L. Sánchez-Pulido, A. Valencia, A. Martínez, Molecular and structural analysis of the panallergen profilin B cell epitopes defined by monoclonal antibodies, *Int. Immunol.* 14 (2002) 993–1001.
- [32] G. López-Torrejón, A. Díaz-Perales, J. Rodríguez, R. Sánchez-Monge, J.F. Crespo, G. Salcedo, L. Pacios, An experimental and modeling-based approach to locate IgE epitopes of plant profilin allergens, *J. Allergy Clin. Immunol.* 119 (2007) 1481–1488.
- [33] A. Vallverdu, J.A. Asturias, M.C. Arilla, N. Gomez-Bayon, A. Martinez, J. Martinez, R. Palacios, Characterization of recombinant *Mercurialis annua* major allergen Mer a 1 (profilin), *J. Allergy Clin. Immunol.* 101 (1998) 363–370.
- [34] S. Laffer, S. Vrtala, M. Duchêne, R. van Ree, D. Kraft, O. Scheiner, R. Valenta, IgE-binding capacity of recombinant timothy grass (*Phleum pratense*) pollen allergens, *J. Allergy Clin. Immunol.* 94 (1994) 88–94.
- [35] M. Sankian, M. Yousefi, N. Pazouki, A. Varasteh, One-step purification of histidine-tagged profilin with high purity and yield by using metal precipitation, *Biotechnol. Appl. Biochem.* 47 (2007) 185–189.
- [36] G. Pauli, J.P. Oster, P. Deviller, S. Heiss, J.C. Bessot, M. Susani, F. Ferreira, D. Kraft, R. Valenta, Skin testing with recombinant allergens rBet v 1 and birch profilin, rBet v 2: diagnostic value for birch pollen and associated allergies, *J. Allergy Clin. Immunol.* 97 (1996) 1100–1109.
- [37] U.K. Laemmli, Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature* 227 (1979) 680–685.
- [38] T. Kleber-Janke, W.M. Becker, Use of modified BL21(DE3) *Escherichia coli* cells for high-level expression of recombinant peanut allergens affected by poor codon usage, *Protein Expr. Purif.* 19 (2000) 419–424.